Animatronic Humanoid Robot System
That Responsively Interacts with Humans

Tomomichi Sugihara*, Kou Yamamoto*, Wataru Takano*, Katsu Yamane*, Yoshihiko Nakamura*
*Department of Mechano-Informatics, Univ. of Tokyo.
7–3–1, Hongo, Bunkyo-ku, Tokyo, 113–8656, Japan.
Email: {sugihara, kou, takano, yamane, nakamura}@ynl.t.u-tokyo.ac.jp

Animatronic humanoid robot system is an integration of robot body with high mobility, information processing
concerning with the motion management of a large-scale dynamical system, and motor control handling with
underactuated articulated body, which are even individually still challenging problems. This paper provides key
technologies towards all those requirements. A reliable miniature robot body is tolerable against large accelerations.
A combination of realtime motion capture and Mimetic Communication model works as a quick decision-making
module. Dynamical motion retouch modifies physically inconsistent referential trajectories to consistent ones only
within one footstep. Finally, they are built into a responsive autonomous humanoid system in a style of virtual
fighting between a real human and the robot.

1. INTRODUCTION

“Animatronics” originally means a technology to realize lively actions of robots, particularly in the field of
movies and entertainments. Now, it has also become a key term for those who develop robots at the stand-
point of industry and academia. Passing by a preliminary study phase of each basic motion such as standing,
walking, crawling, rising-up, etc., how to synthesize animated behaviors of robots from a variety of motion
primitives is being focused on.

Aiming at imitating not only apparent animations but the internal mechanism of humans, this problem should
be approached from both sides of hardware and software. The former is about the design of robot bodies
with high reliability and potentially high mobility. The latter includes widely-ranged fields, such as measure-
ment of the external world to reconstruct them inside of computers, information processing to manage
humanoid action, and motor control to achieve actual robot motions. Furthermore, those functions should be
interlocked and cycled rather at high rate for human-like lively robots that can cope with the changeful world
every moment. This requirement makes the problem extremely difficult even as individual issues. Although
some sophisticated humanoid robots [1] [2] [3] [4] [5] [6] that are almost autonomous with the capabilities of
recognizing objects and events outside the body, planning the action according to those information and output-
ning them with real bodies have already been presented, they basically stand on the conventional logic-based
AI's and a path-planning with sufficiently long time to plan.

Fig.1 Outer view of UT-μ2

This paper provides the key technologies for animatronics, which cover i) a small anthropomorphic robot, ii) motion capturing system to analyze human behaviors in realtime, iii) brain-like information processing system for decision making, and iv) motion retouch to embody the decisions as robot actions. They are integrated into a primary but completely autonomous humanoid robot system. In order to show the power of them, a virtual fighting between a real human and the robot is demonstrated as an example of a quick and dynamic interaction.

2. A MINIATURE HUMANOID ROBOT

2.1 Scope

One of the issues for animatronic humanoid control is a requirement of large accelerations involved by dy-
namic responsive motions with frequent collision and
detachment with the environment. From this point of view, the use of a light miniature body is preferable in this starting stage of research. A small body, however, causes some constraints on the assembly. Self-collision, for instance, could happen easily. And, the space to mount mechatronic components and to wire the cables is restricted. For reliable bodies, it is needed to devise the design.

There has been several attempts to build small lightweight humanoid robots. Remote-brained robot approach proposed by Inaba et al. [7] released the robot body from heavy computers. Their bodies consist of tens of small servo DC motor modules originally for radio-controlled toys, and have evolved to humanoid types [8]. Using those series, Nagasaka et al. [9] realized various whole-body motions. On the same technology, Nordin et al. [10], Furuta et al. [11] and Yamasaki et al. [12] also developed low-cost anthropomorphic robots. However, low precision on the control of those actuators prevents reliable dynamic motion controllers to be implemented. Kuroki et al. [13] developed SDR-4X. And, Fujitsu Automation [14] produced HOAP series for the purpose of offering an open platform to researchers. They developed unique components including motors, gears, amplifiers, I/O interfaces and so forth to be integrated in them. A large part of those technology is not revealed. On the other hand, small types of mechatronic products with high specification to be embedded especially into robots have recently become commercially available. Common developers have gained the choice of them for their own systems. This section introduces a miniature humanoid robot designed with such public components to overcome the technical hurdles. Kinematic issues including the motion range are firstly dealt with modularized ortho-axis-coupled joint units, and the issue of integration and wiring is reduced by AnimatoCore.

2.2 Specifications

Fig.1 shows the robot we developed and named UT-μ2. It is about 54[cm] tall, and weighs about 7.5[kg] including the battery and processor on the body. The total number of joints is 23; 3 for the neck are without actuation. 4 located at each arm, and 6 at each leg are actuated by coreless DC motors manufactured by maxon motor ag., and decelerated with harmonic drive gears by Harmonic Drive Systems. The gear ratios are all 100:1, and the output of motors are 4.5W (for elbow rotational joints), 6.5W (for shoulder abdutional joints, elbow flexional joints and knee rotational joints), 22W (for hip abductional joints) and 11W (for the other joints), respectively. The choice of those high specifications mounted on the arms and legs encourages aggressive types of the whole body motion, and also promises feeding the results of experiments back to life-sized robots. Although it causes an increase of weight at each joint, the thin-shelled exoskeletal structure, made of magnesium alloy casting, is designed as curved surface for high stiffness versus weight, so that the total weight meets the standard weight-per-size ratio. In addition, the total number of parts is reduced, and much higher unit cost of magnesium alloy casting than milling and bonding is balanced.

2.3 Ortho-axis-coupled joint unit

![Fig.2 Joint assignment of UT-μ2](image)

![Fig.3 Ortho-axis-coupled joint unit](image)

Each joint is modularized as a unit with coupled axes, named ortho-axis-coupled joint unit. The idea is basically in accordance with the fact that many of the couples of adjacent axes in anthropomorphic systems are mutually-perpendicular. The whole-body of the robot is assembled, connecting those units with magnesium exoskeletal structures. Fig.3(A) illustrates the unit and the structural connectivity around it. Each axis involves a pair of gear and motor. Fig.3(B) is the picture of the gear part. Adopting this technique, 1) the robot keeps maintainability, getting attachment-and-detachment of each joint to be easy, and 2) the production cycle is accelerated, sharing common parts. Fig.2 shows the total joint assignment of the robot. All the units are located to avoid easy self-collision. Some unique features in it are as follows. Two axes of the hip
joints are assigned so as to enlarge the motion range of flexion-extension and abduction. The distance between the two hip joints is countered by inward offset of knee rotation joints to prevent the sideward perturbation of the trunk. The ankle flexional joint is assigned with 5 mm offset backward to the hip and knee in order to exclude singular posture out of the motion range. And, the mechanical stopper at the knee flexional joint saves kinematic energy in the stretched-knee posture. The root axis of each shoulder is slanted from the perpendicular axis to 45 degrees, aiming at a natural flavor on its mixed motion of extension and abduction in a simple way.

2.4 AnimatoCore

The electronic components only except actuators and sensors of the robot are integrated to be a core control unit, independently from the robot body. We named it AnimatoCore. Fig.4 is its illustration and outer view. This technique largely changes the production cycle from that of conventional humanoid robots, which contain both mechanical and electrical components in their own frames. Namely, since the development of electrical system is separated from that of mechanical system, repetitive unmount-and-remount operation of components especially in the early stage of the development becomes less troublesome. And, cable wiring in the limited space inside of the body is simplified, so that the risk of self-destruction around mechanically complex parts such as wire-breaking is reduced. In addition, fragile electric circuits are excluded from the extremities, which frequently suffer from vibration and perturbation during the motion. Moreover, it could function as a stand-alone portable controller of various space-saving robotic systems such as mobile robots, multi-legged robots, surgical robots and so forth. One constraint is that this architecture conditions mechanical design to have sizable space to mount the unit.

Fig.5 is a diagram of the electronics system of UT-μ2 centering around AnimatoCore. One can see that it is independent in the sense that it has a processor board and battery within, excluding any cables from outer resources which disturbs robot’s motion. The processor board is E/Kit-1100 (Device Drivers) featuring A/D converter and two CF slots, one of which is for wireless LAN (corega) and the other for storage. We also developed a small device controller board with H8 processor (Renesas Technology Corp.), two-channel DC motor drivers and eight-channel A/D converters on the cooperation with STEC Corp. Fig.6 shows the external view of the board. The internal network between the main processor and distributed motor/sensor controllers on the board is built with CUnet (StepTechnica), which realizes widely-ranged multi-CPU network on shared memory.
Fig. 7 Diagram of virtual fighting system

The input signal from sensors, which includes the gyro sensor CRS-03 (Silicon Sensing Systems Ltd.) the accelerometer G5X010011T (Matsuky Co., Ltd.) and the 3-axis force sensor PicoForce (NITTA), are processed at A/D converter on the device board. 3-axis force sensors are mounted on the soles and wrists, 4 for each sole and 1 for each wrist. Those on the sole are located at four corners to calculate the resultant linear force and moment applied to each foot from the ground.

3. RESPONSIVE HUMAN-ROBOT INTERACTION SYSTEM

3.1 System overview

This section shows an architecture of an autonomous humanoid system which responsively interacts with humans through observation and action. It forms a style of virtual fighting between a real human and a robot. Since fighting is a quick type of communication, it's worth mentioning that observation-action cycle is so short that it's beyond the performance of a combination of conventional logic-based intelligence and offline motion planning. Much higher system responsibility is required, and then, we adopted Mimesis Loop paradigm[15] extent to a realtime implementation for the brain part.

Fig.7 and Fig.8 show the system construction and the actual scene of the fighting field, respectively. The system consists of three parts, namely, i) observer-tactician part including optical motion capture system to observe player's action and Mimesic Communication model[16] to make decisions about tactics, ii) robot motion controller part including online dynamical motion planner and iii) image synthesizer part which recomposes virtual fights on the display.

The online motion recomposition method shown in the previous section works as a bridge between a higher-level information processing module as an autonomous decision-making mechanism and a lower-level robot motion controller. In other words, it could complement the deep valley between the artificial intelligence and physically embodied robot system.

The player's motion is captured and converted to the motion of a human-figure which has the same joint assignment with the robot. Then, Mimesic Communication model recognizes it, selects the corresponding motion as a tactics and sends the sets of the desired posture and supporting state with a period of 30[ms] to the robot. The PC specification on which system runs is CPU: Xeon 3.60GHz × 2 and 2GB RAM.

Robot motion controller receives them via TCP/IP using socket connection, and interpolates it into dynamically consistent referential motion every 3[ms].

The PC specification on which system runs is CPU: Pentium 4 3.00GHz and 512MB RAM. And, as is described in the subsection 2.4, the robot itself has a processor to control motors and to process sensor signals. The main PC and it communicates with each other via wired TCP/IP. The robot status is also monitored every 3[ms] and sent to the above Mimesic Communication system.

Human-figure motion and the real robot motion monitored are virtually synthesized on another PC with CPU: Xeon 3.60GHz, 2GB RAM and nVIDIA FX 4400, and the image of virtual ring and two fighters is displayed.

In this demonstration, average time of supporting state transition, which also means the average latency of communication between decision-making part and motor control part, is about hundreds of milliseconds.
3.2 Realtime motion capture

MAC3D System (Motion Analysis Corporation) is adopted for motion capture system. It measures the positions of retroreflective markers attached to the actor’s body, and labels all of them based on the stored structure of human model.

The posture of human-figure in the global frame is calculated from those labelled points by inverse kinematics. This task involves some technical difficulties such as singularity and redundancy due to the complexity of a human model. A large number of links and joints also causes a practical issue of computation cost, which realtime implementation suffers from. Then, we applied the method proposed by Yamane et al. [17]

3.3 Mimetic communication [16]

Mimesis Loop explored by Inamura et al. [15] is a framework in which intelligence to acquire and invoke protosymbols is successively cultivated through the process of imitating others’ behaviors, namely, recognizing them, remapping them into the robot’s body and estimating the likelihood between the original motion observed and that replayed by the robot, using Hidden Markov Model (HMM). Based on this Mimesis criterion, Takano et al. [16] proposed a primitive communication model which consists of hierarchical HMM. In the learning phase, the lower layer memorizes motion patterns of two persons independently, outputs the identifier of the HMM storing the pattern and gives it to the upper layer. The upper layer memorizes the interaction patterns, namely, the combination of the motion identifier of those by two individuals. Eventually, the model acquires the correspondence of those behaviors which is thought to be resulted of communication between them.

The most original property of the model lies on that it also functions as an online behavioral decision system of the robot. From the history of interaction pattern between the robot and the other with one additional lat-
est action of the other recognized, the upper layer finds HMM which stores the most likely pattern to the information with a partial lack, chooses the motion identifier which corresponds to the missing part and decides it as robot’s next motion to appear. Fig.9 illustrates the mechanism of Takano’s model. It is utilized in our system as a tactician module; picking up the “will” of the player whose action is measured by the motion capture system, determining abstract tactics such as “punching”, “avoiding” and so forth, choosing the representative sequence of the desired posture and supporting state with a period of 30[ms], and sending them to the robot as a motion command. The posture includes the whole joint angle values and the attitude of trunk link with respect to the inertial frame.

3.4 Dynamical motion retouch using Boundary Condition Relaxation [18]

Physical constraints to be noted when designing referential motion trajectory for humanoid robots are itemized as follows.

1. Let joint angles be within the motion range
2. Avoid self-collision
3. Smoothen the trajectory so as to be apparently continuous
4. Not require attracting force acting at each contact point to the environment
5. Follow the planned supporting state transition
6. Carry ZMP [19] onto the sole of supporting foot before detaching the swing foot off the ground
7. Let joint angle velocity be within the limitation
8. Let joint torques be within the limitation

1)~3) are in the group of geometric constraints, while 4)~6) are in the group of dynamical constraints. And, 7)8) are concerned with joint actuator specification. 4), the severest constraint, is attributable to the fact that humanoid robots lack of mechanical connection to the inertial frame, and is substitutable for a condition that ZMP locus of the motion stays within the planned supporting region. 5) and 6) also involve it. Particularly, these requirements about dynamics condition indicate that not only desired postural sequence but also supporting state transition command is needed to describe humanoid motion with respect to the environment. Although there has been proposed several techniques to modify a trajectory with physical inconsistency to be applied for real robots [20] [21] [22] [23] [24], they all work in offline, and in addition, some of them don’t deal with the above severe dynamical constraint.

Responding to the above indication, we suppose that a higher-level information processing module steadily outputs a rough transition sequence of desired posture and supporting condition with a period of tens of millisecond as a motion command. In intuitive sense, this
corresponds to the motion image we humans associate with symbolic motion identifier in mind. And, dynamics is not taken into account in this stage, namely, the constraints 4) and 6) in the previous subsection are not necessarily satisfied. For simplicity, however, we assume that the original rough sequence already satisfies 1)2)7)8) and has enough margin, so that trajectory modification in the following subsection doesn’t harm those constraints, and also satisfies 3)5).

Under those assumptions, one can segment the input sequence according to the transient of desired supporting condition as is figured by Fig.10. Once the changeover of the desired supporting condition is recognized, a sequence of contact state transition motion is clipped, and the following online trajectory modification is applied for it. Note that it causes a latency to wait the end of the segment between the time when the beginning of it is given and when it is referred from the real robot in principle.

Suppose a segmented subsequence is given, and let $T$ be its time length. The online dynamical trajectory modification is processed as follows.

1) Create cubic spline curves which connect given desired via-points of the extremities and trunk attitude. Let us suppose it is $\mathbf{d}_s(t)$, here. The problem is that, even if the original $\mathbf{d}_p(t)$ is continuous and smooth, the actual robot state does not necessarily coincide with the desired at the end of each segment and discontinuity occurs, due to model error, the effect of gravity, computational error in inverse kinematics and so forth. Then, as is shown in Fig.11(A), supplement function $\Delta \mathbf{p}(t)$, defined by (2) for example, to the original $\mathbf{d}_p(t)$ (chained line in the figure) in order to reduce the gap between the desired previous goal, which is the same with the desired initial value of the new segment, and the actual initial value to be zero, taking time $T$. The resultant referential trajectory $\mathbf{p}(t)$ concatenated to the actual initial robot state forms as follows.

$$
\mathbf{p}(t) = \mathbf{d}_s(t) + \Delta \mathbf{p}(t)
$$

$$
\Delta \mathbf{p}(t) \equiv (\mathbf{p}(0) - \mathbf{d}_s(0)) \left(1 - \frac{t}{T} + \sin 2\pi \frac{t}{T}\right)
$$

II) Compute the referential trajectory of the center of gravity (COG) from the initial state (position and velocity) and the desired goal state after $T$, using Boundary Condition Relaxation[18] as is figured in Fig.11(B). $\mathbf{p}_c(t)$ is gained as an analytical function based on an approximate dynamical model whose mass is concentrated at COG. The significance of the adoption of this method is that it creates referential COG trajectory $\mathbf{p}_c(t)$, and simultaneously ZMP trajectory $\mathbf{p}_z(t)$, only from the next desired goal state and supporting condition transition completely online, admitting some error from the desired state at the end of the segment. Though it is pragmatic in the sense that it ignores the moment around COG derived from distributed mass and inertia of each link, it can rapidly compute a set of COG/ZMP trajectory which satisfy the dynamical constraint 4).

III) In the previous stage, the time of foot-detachment $T_D$ is found from the planned ZMP trajectory $\mathbf{p}_z(t)$ and the desired supporting region $\mathbf{d}_S$; $T_D$ is the minimum time when $\mathbf{p}_z(t)$ comes within $\mathbf{d}_S$. In order to ensure the consistency of ZMP trajectory and foot placement, the foot to move should keep contact with the ground until $T_D$. Then, as is figured in Fig.11(C), timesteps $p(t)$ originally defined in $0 \sim T$ to $p'(t)$ in $T_D \sim T$ as follows.

$$
p'(t) = \begin{cases} 
p(0) & (0 \leq t < T_D) 
\frac{p(T(t-T_D))}{T-T_D} & (T_D \leq t < T) 
\end{cases}
$$

As the result of the above procedure, modified trajectories of COG, trunk attitude and the extremities which satisfy the conditions 1)–8) are generated. Note that all those trajectories are designed as continuous functions, so that it is quantized-period invariant. Sampling the set of referential values from them and applying inverse kinematics of the whole body [17], the referential joint angle values are obtained.

4. DEMONSTRATION

Fig.12 shows a sight of the actual performance of a virtual fighting between a real human and a robot. It is an example of a quick and dynamic interaction which could hardly be realized by conventional autonomous system.

This system does not feature a bilateral force feedback mechanism, and collision detection is done only by geometric computation between the models. Players attack the imaginary opponent, watching the synthesized image in the display. In the sense that the information to the player was restricted only to visual sensation, it lacked of reality as a virtual type of application. However, the initial objective that a robot autonomously decides its behavioral strategy, recomposes it to be physi-
cally consistent motion trajectory and performs according to them was achieved.

5. CONCLUSION

We presented hardware and software approaches towards animatronic control of humanoid robots. In order to show how the total system works, a virtual fighting between a real human and the robot were demonstrated.

Technical difficulties to build reliable miniature humanoid robots range from kinematical ones to those in the integration. We proposed joint modularization with ortho-axis-coupled joint unit and easy construction of the robot body with them, and an electronic system architecture centering around the core control unit named AnimatoCore to ease the troubles with integration and wiring in the limited space, through the design of UT-μ2. The use of miniature types of robots are advantageous for quick motions which necessitates large accelerations.

From the viewpoint of the system architecture, new paradigms were required on both brain-part and motor controller which cyclicly worked at high rate. A combination of realtime motion capture and Mimetic Communication model provides such a responsive intelligent system. And, Boundary Condition Relaxation which is based on a simplified dynamics model realizes the online motion retouch which converts a physically inconsistent state transition sequence to a consistent one only within one step, though it has hundreds of milliseconds of latency.

Acknowledgments

This research was supported partly by Project for the Practical Application of Next-Generation Robots in The 21st Century Robot Challenge Program, New Energy and Industrial Technology Development Organization, and partly by Category “S” of Grant-in-Aid for Scientific Research, Japan Society for the Promotion of Science.

The authors greatly appreciate to NAC Image Technology, Inc. and Motion Analysis Corporation for letting us use the MAC3D System. And also thank Mr. Tatsuhito Aono as a test player of the virtual system demonstrated in section 3.
References


